PRIOR MEASURE FOR NONEXTENSIVE ENTROPY

F.Brouers and O.Sotolongo-Costa

Institute of Physics, Liège University and Henri Poincaré Chair of Complex Systems, Havana University, Cuba.
fbrouers@ulg.ac.be , oscarso@ff.oc.uh.cu

ABSTRACT

We show that if one uses the invariant form of the Boltzmann-Shannon continuous entropy, it is possible to obtain the generalized Pareto-Tsallis density function, using an appropriate "prior" measure $\mathrm{m}_q(\mathrm{x})$ and a "Boltzman constraint" which formally is equivalent to the Tsallis q-average constraint on the random variable X. We derive the Tsallis prior function and study its scaling asymptotic behavior. When the entropic index q tends to 1, $\mathrm{m}_q(\mathrm{x})$ tends to 1 for all values of x as this should be.

PACS: 0590+m, 02.50.-r, O5.45.Df

1 Introduction

Most of the probability distributions used in natural, biological, social and economic sciences can be formally derived by maximizing the entropy with adequate constraints ($maxS$ principle)[1].

According to the $maxS$ principle, given some partial information about a random variable *i.e.* the knowledge of related macroscopic measurable quantities (macroscopic observables), one should choose for it the probability distribution that is consistent with that information but has otherwise a maximum uncertainty. In usual thermodynamics, the temperature is a macroscopic observable and the distribution functions are exponentials.

Quite generally, one maximizes the Shannon-Boltzmann (S-B) entropy:

$$S_B = -\int_a^b p(x) \ln p(x) dx \tag{1}$$

subject to the conditions

$$\int_a^b p(x)dx = 1, \int_a^b g_i(x)p(x)dx =< g_i(x) >, \quad i = 1, 2, .. \tag{2}$$

Both limits $a$and b may be finite or infinite. The functions $g_i(x)$whose expectation value have been usually considered [1] as constraints to build probability distributions are of the type

$$x, x^2, x^n, (x- < x >)^2, | x |, | x- < x >|, \ln x, (\ln x)^2, \ln(1 \pm x), \exp(-x).... \tag{3}$$

The maximum entropy probability density function (*mepdf*) depends on the choice of the limits of integration a and b and the functions $g_i(x)$ whose expectation values are prescribed.

One constructs the Lagrangian

$$L = -\int_a^b p(x) \ln(p(x)dx \tag{4}$$

$$-\lambda_0(\int_a^b p(x)dx - 1) - \sum_i \lambda_i(\int_a^b g_i(x)p(x)dx - < g_i(x) >$$

and differentiating with respect to $f(x)$, one finds easily :

$$p(x) = C \exp[-\sum_i \lambda_i g_i(x)] \tag{5}$$

The factor C is a normalization constant and the Lagrange parameters λ_i are determined by using the constraints (eqs.2,3).When this cannot be achieved simply, the parameters λ_i are defined by the constraints. Most distributions derived from the constraints given in eq.3 posses finite second moments and hence belong to the domain of attraction of the normal distributions. Those which belong to the domain of attraction of the Lèvy (stable) distribution *i.e.* the Cauchy and the Pareto distributions are obtained with a characteristic Lèvy tail parameter $\mu \geq 1$ indicating that only a finite expectation value (first moment) can be defined. In particular using the simplest constraint

$$\int_0^\infty xp(x)dx =< x > \tag{6}$$

one obtains the density

$$p(x) = (1/ < x >) \exp(-x/ < x >) \tag{7}$$

which is the basis of Boltzmann thermostatistics if we identify x with the energy E and define the temperature as $T = k < E >$ where k is the Boltzmann constant.

2 Shortcomings of the $S\max$ principle

This traditional utilization of $maxS$ principle has two main shortcomings.

The first one is that it cannot provides distributions belonging to the basin of attraction of Levy distributions when the Levy heavy tail index $\mu < 1.i.e.$distributions corresponding to the statistial properties of non-averaging systems. Applied to thermostatistics this means that it cannot be applied to nonextensive systems.

The second, pointed out repeatedly for decades by Jaynes[2][3], is that the continuous form of the Boltzman-Shannon entropy is not invariant and should be written more correctly as

$$S = -\int p(x) \ln \frac{p(x)}{m(x)} dx \tag{8}$$

in order to have a distribution invariant under parameter change. The function $m(x)$ is called the measure function. It is proportional to the limiting density of discrete points. So it is the entropy, relative to some measure, which has to be maximized.

Under a change of variable the function $p(x)$ and $m(x)$ transform in the same way. If we maximize the entropy

$$\int p(x) \ln \frac{p(x)}{m(x)} dx \tag{9}$$

with the constraint

$$\int g(x) p(x) dx = 1 \tag{10}$$

The Lagrange multiplier method yields for $p(x)$ the solution

$$p(x) = Cm(x) e^{-\lambda g(x)} \tag{11}$$

The meaning of the measure function has been discussed at large by Jaynes [2][3]. If there is no constraint maximizing the entropy yields $p(x) = Cm(x)$ where C is a normalizing constant. It is the distribution representing "complete ignorance"

3 Tsallis entropy

The most popular method to circumvent the first difficulty is based on a generalization of the Boltzmann entropy known as the Tsallis entropy[4][6]. This method has been used with some success to deal with thermostatistic properties of slightly chaotic and nonergodic systems presenting memory effects and long-range interactions.

The Tsallis entropy is appropriate for system weakly chaotic power law mixing when the Liapounov exponent is tending to zero and the relevant phase space instead of being translational invariant (as in the Gibbs-Boltzmann case) is scale invariant (fractal or more generally multifractal). We start from the Tsallis nonextensive entropy

$$S_T = -\int_0^\infty p_q^q(x) \ln_q p_q(x) dx \tag{12}$$

subject to the conditions

$$\int_0^\infty p_q(x)dx = 1, \int_0^\infty g_{q,i}(x)\tilde{p}_q(x)dx = < g_{q,i}(x) >_q, \quad i = 1, 2, .. \tag{13}$$

$$\tilde{p}_q(x) = \frac{p_q^q(x)}{\int_a^b p_q^q(x)dx} \quad f_q(x) = \frac{\tilde{p}_q^{1/q}(x)}{\int_a^b \tilde{p}_q^{1/q}(x)dx} \tag{14}$$

where the functions $\tilde{p}_q(x)(x)$ are the so-called "escort" probability [6][5] . Generalizing what has been done with the standard S-B entropy (eqs.1-4), one can write the constraints using the generalized q-exponential and q-logarithm functions (q-constraints).

$$\ln_q x \equiv \frac{x^{1-q} - 1}{1-q} \qquad \exp_q(x) \equiv [1 + (1-q)x]^{\frac{1}{1-q}} \quad \text{with} \quad \ln_q(\exp_q(x)) = x \tag{15}$$

construct the corresponding Lagrangian and differentiating with respect to $\tilde{p}_q(x)$ one finds [8]

$$p_q(x) = C_q \exp_q[-\sum_i \lambda_{q,i} g_{q,i}(x)] \tag{16}$$

$$\tilde{p}_q(x) = \tilde{C}_q p_q(x)^q \tag{17}$$

The factors C_q are normalization constants and the Lagrange parameters $\lambda_{q,i}$ are determined by using the q-constraints. If one chooses the simplest constraint with $g_{q,i}(x) = x$ defining the so- called q-average

$$\mathrm{E}_q(x) = \int_0^\infty x)\tilde{p}_q(x)(x)dx = < x >_q \tag{18}$$

we get the generalized Pareto density supported on the positive half-line and defined in the range $1 < q < 2$.

$$\tilde{p}_q(x) = (1/(2-q) < x >_q)[1 + \frac{q-1}{2-q} \frac{x}{< x >_q}]^{-\frac{q}{q-1}} \tag{19}$$

$$p_q(x) = (1/ < x >_q)[1 + \frac{q-1}{2-q} \frac{x}{< x >_q}]^{-\frac{1}{q-1}} \tag{20}$$

and the corresponding Pareto-Tsallis distribution supported on the positive half-line.

$$P_q(x) = 1 - [1 + \frac{q-1}{2-q} \frac{x}{< x >_q}]^{-\frac{q}{q-1}} \tag{21}$$

Defining the tail index $\mu = \frac{2-q}{q-1}$, the distribution $P_q(x)$ has an asymptotic algebraic behavior

$$p_q(x- > \infty)- > (1/\mu)\ x^{-\mu-1} \tag{22}$$

$$P_q(x- > \infty)- > (1/\mu)\ x^{-\mu} \tag{23}$$

The n-th moments E[x^n] of the random variable of the random variable X distributed according to eq.() exists only if $n{<}\mu$ (hence for $1 < q < \frac{2+n}{n+1}$).Therefore $P_q(x)$ belongs to the domain of attraction of the completely assymmetric stable Lévy distribution. These properties and their consequences are widely discussed in the literature.[5][7]

4 Determination of the prior for nonextensive entropy

We can now come to the second shortcoming. The meaning of the measure function has been discussed at large by Jaynes[2][3]. In the continuous case even before we can apply maximum entropy principle, we must deal with the problem of complete ignorance (for instance all microstates are equiprobable in Boltzmann-Gibbs thermodynamic). For "fractal" systems ignorance means scaling invariance. It is therefore legitimate to ask what is the scale invariant measure function or "prior" which should be used to obtain the Generalized Pareto-Tsallis density function if we would use the correct invariant form

$$S_I = -\int_0^\infty p_q(x) \ln \frac{p_q(x)}{m_q(x)} dx \tag{24}$$

and use the constraint

$$< g(x) >= \int g(x) p_q(x) dx \tag{25}$$

Starting from the identity

$$S_I = -\int p_q(x) \ln \frac{p_q(x)}{m(x)} dx = S_T = \int (p_q(x))^q \ln_q p(x) dx \tag{26}$$

we derive the following relations:

$$\ln p_q(x) - \ln m(x) = (p_q(x))^{q-1} \ln_q p_q(x) \tag{27}$$

$$S_T = S_B + < \ln m(x) >$$

or defining the "Kullback-Leibler divergence" $KL(p(x), g(x))$

$$KL(p_q(x), m(x)) = S(p_q, m) - S_B \tag{28}$$

the quantity

$$S(p_q, m) = - < \ln m(x) > \tag{29}$$

being known as "cross-entropy". The cross entropy is always greater than or equal to the entropy, this shows that the Kullback-Leibler (KL) divergence is always nonnegative and furthermore $KL(p(x), p(x))$ is zero. We obtain easily

$$\frac{p_q(x)}{m_q(x)} = \exp[(p(x))^{q-1} \ln_q p_q(x)] \tag{30}$$

We use the definition of $\ln_q$ (eq.15), and obtain

$$m_q(x) \ = p_q(x) \exp(\ln_q(\frac{1}{p_q(x)}) \tag{31}$$

If we choose for $p_q(x)$ the generalized Pareto-Tsallis function with reduced variable ($< x >_q = 1$)

$$p_q(x) = (1 + \frac{q-1}{2-q} x)^{-\frac{1}{q-1}} \tag{32}$$

We obtain easily the following results:

$$\ln_q(\frac{1}{p_q(x)}) = \frac{x}{(2-q) + (q-1)x} \tag{33}$$

$$m_q(x) = \exp_q(-\frac{1}{2-q} x) \exp(\frac{x}{(2-q) + (q-1)x}) \tag{34}$$

If we define the $q-$Laplace transform in he sense of Lenzi et al. [9] :

$$\mathcal{L}_q[f(t)](s) \equiv F_q(s) \equiv \int_0^\infty f(t)[\exp_q(-t)]^s dt \tag{35}$$

we have

$$\int_0^\infty m_q(x) dx = F_q(1) = \int_0^\infty f(t)[\exp_q(-t)] dt \quad \text{with } f(t) = \frac{t}{1 + (q-1)t} \tag{36}$$

In term of the Levy tail index

$$\mu = \frac{2-q}{q-1} \tag{37}$$

we can write

$$m_q(x) = (1 + \frac{1}{\mu}x)^{-\mu-1})\exp(\frac{(\mu+1)x}{\mu+\ x}) \tag{38}$$

This function obeys the following asymptotic behaviors

$$m_q(x->\infty)->\exp(1+\mu)\frac{1}{\mu}x^{-\mu-1} \tag{39}$$

$$m_q(x->0) = 1-(\mu+1)(x/\mu)^2+... \tag{40}$$

The prior function $m_q(x)$ has the scaling invariant asymptotic form of a completely asymmetric Lévy stable distribution. In Fig.1 and Fig.2 we represent the function for two value of q (1.2 and 1.8) coresponding to values of the tail index μ (4 and 0.25) and for the physical range $1 < q < 2$.

5 Maximization of the invariant entropy

If we maximize the entropy

$$S_I = \int p_q(x)\ln\frac{p_q(x)}{m(x)}dx \tag{41}$$

with the constraints

$$\int_0^\infty p_q(x)dx = 1 \qquad \int_0^\infty g(x)p_q(x)dx = 1 \tag{42}$$

The Lagrange multiplier method yields for $p(x)$ the solution

$$p_q(x) = Cm_q(x)e^{-\lambda g(x)} \tag{43}$$

We will recover the Tsallis function

$$p_q(x) = (1+\frac{q-1}{2-q}x)^{-\frac{1}{q-1}} \tag{44}$$

if

$$\frac{x}{(2-q)+(q-1)x} = (1/(2-q))\frac{x}{1+(q-1)/(2-q)x} = \lambda g(x) \tag{45}$$

i.e.

$$\lambda = 1\ ,\ (<g(x)>=1) \text{ and } g(x) = (1/(2-q))\frac{x}{1+(q-1)/(2-q)x} \tag{46}$$

In that case the constraint (equation) reads

$$\int g(x)p_q(x)dx = (1/(2-q))\int x(1+\frac{q-1}{2-q}x)^{-\frac{q}{q-1}}dx = 1 \tag{47}$$

which is the Tsallis q-average constraint on the variable x since

$$\tilde{p}_q(x) = (1/(2-q))(1 + \frac{q-1}{2-q}x)^{-\frac{q}{q-1}} \tag{48}$$

is the Tsallis escort probability (eq.19). Therefore the use of the correct non invariant form

$$S_I = -\int_0^{\infty} p_q(x) \ln \frac{p_q(x)}{m_q(x)} dx \tag{49}$$

with the "prior"

$$m_q(x) = \exp_q(-\frac{1}{2-q}x) \exp(\frac{x}{(2-q)+(q-1)x}) \tag{50}$$

and the constraint

$$\int g(x) p_q(x) dx = 1 \tag{51}$$

with

$$g(x) = (1/(2-q)) \frac{x}{1+(q-1)/(2-q)x} \tag{52}$$

yields the generalized Pareto density

$$p_q(x) = (1 + \frac{q-1}{2-q}x)^{-\frac{1}{q-1}} \tag{53}$$

and the "Boltzmann" constraints (eq.) is equivalent to the Tsallis constraints $< x >_q = 1$.

6 Conclusions

We have shown that maximizing the invariant continuous Boltzmann-Shannon entropy with appropriate prior measure and constraint provides the Generalized Pareto Tsallis distribution which is the basis of the nonextensive thermostatistic. This point of view opens paths for deriving superstatistics directly from the "universal" invariant Boltzmann-Shannon entropy.

7 Acknowledgment

One of us is grateful to Prof.N.Kumar for useful discussions.

References

[1] Kapur J.N , Maximum Entropy Models in Science and Engineering , Wiley and Sons, New York (1989).

[2] Jaynes E.T.,Prior Probabilities IEEE, Transaction on System Science and Cybernetics SSSC-4 (1968), 227.

[3] Jaynes E.T, Probability Theory, The Logic of Sciences, Cambridge University Press (2003).

[4] Tsallis C., J.of Stat.Phys. 52 (1988) 479.

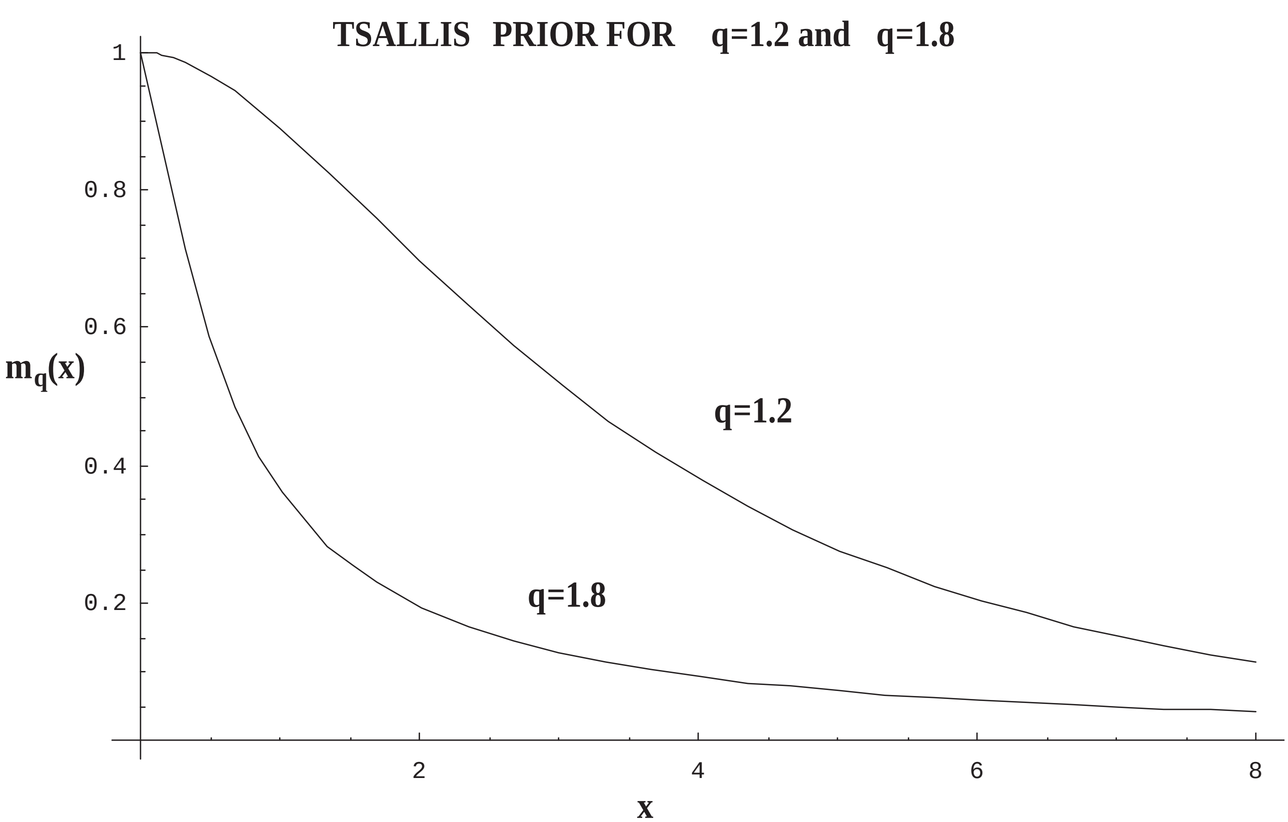
[5] Tsallis C. in Non extensive statistical mechanics and thermodynamics (Ed.S.Abe and Y.Okamoto, Springer Ed., Berlin 2001).

[6] Tsallis C.,Mendes R.S.,and Plastino A.R. PhysicaA A261 (1998) 534.

[7] Abe S. J.Phys A Math.Gen. 36 (2003) 8733.

[8] F.Brouers, O.Sotolongo-Costa and K.Weron, PhysicaA, 344/3-4 (2004) 409

[9] Lenzi E K,Borges E P and Mendes S R J.Phys A Math. Gen. 32 (1999 8561.


TSALLIS PRIOR FOR q=1.2 and q=1.8
1
0.8
0.6
0.4
0.2
$m_q(x)$
q=1.2
q=1.8
2
4
6
8
x

TSALLIS PRIOR